The Interventionist Account of Causation and Non-causal Association Laws

Max Kistler
Université Paris 1
mkistler@univ-paris1.fr

Erkenntnis, March 2013, DOI **10.1007/s10670-013-9437-4.**
http://link.springer.com/article/10.1007/s10670-013-9437-4

Abstract

The key idea of the interventionist account of causation is that a variable A causes a variable B if and only if B would change if A were manipulated in the appropriate way. This paper raises two problems for Woodward's (2003) version of interventionism. The first is that the conditions it imposes are not sufficient for causation, because these conditions are also satisfied by non-causal relations of nomological dependence expressed in association laws. Such laws ground a relation of mutual manipulability that is incompatible with the asymmetry of causation. Several ways of defending the interventionist account are examined and found unsatisfying. The second problem is that it often seems to be impossible, in a model that contains variables linked by an association law, to satisfy the conditions imposed on interventions on such variables. Various ways to solve this second problem, most importantly the analysis of manipulability in terms of difference making, are examined. Given that none solves the problem, I conclude that the interventionist conditions are neither sufficient nor necessary for causation. It is suggested that they provide an analysis of nomological dependence, which may be supplemented with the notion of a causal process to yield an analysis of causation.

1. The interventionist analysis of causation

According to the interventionist account, causation is a relation between variables. Its fundamental hypothesis is that a variable A causes a variable B if and only if it there are circumstances in which it is possible to manipulate B by intervening on A. According to Woodward (2003; 2008), this idea underlies scientific research for causes across all sciences. He gives the following example from social science. One can observe, in the contemporary US, a statistical correlation between children's attendance of private schools (P) and their scholastic achievements (A). A randomized experiment would be a straightforward way by which a social scientist could try to find out whether this correlation stems from the fact that attendance of private schools causes better scholastic achievement or whether both variables are effects of some common cause, such as the parents' higher socio-economic status (S). Such an experiment requires attributing children from a group of fixed S randomly to two sub-groups: one sub-group is sent to a public school, the other to a private school. This is equivalent to attributing one value of P to the individuals in the experimental group and another value to those in the control group. Making the attribution to the two subgroups random is intended to make it independent of any other factors that could influence A independently from P. After a suitable lapse of time, A is measured in the two subgroups. Any correlation that is found between A and P can be taken to reflect the existence of a causal influence of P on A. The possibility that A and P be the effects of some common cause such as S has been excluded by randomizing the attribution of a value to variable P for each individual. This is supposed to ensure that P is statistically uncorrelated with S, and indeed with any other variables that might be common causes of P and A.

The interventionist account of causation is even more plausible as an analysis of causation in experimental sciences such as physics or chemistry, where it is often practically easier to control the values of variables than in the social sciences. It is rather more difficult to arbitrarily set the value of socio-economic status for a given individual, than to set the value of variables such as the intensity of electric current in a copper wire in a physics laboratory. Let S, A and P represent physical variables characterizing observable and manipulable properties of copper wires. Let P be the electric current flowing through the wire, A the heat release from the wire, and S the room temperature. The fundamental idea of the interventionist account is that P causes A if and only if the following is true: if the room temperature S, as well as all other variables that might influence A are held fixed (except of course P and A themselves), then, if one intervened on P by changing its value, without directly intervening on A or other potential causes of A such as S, then the value of A would change.

Causal relations between variables can be represented by *graphs*. A graph is a pair whose constituents are 1) a set V of variables and 2) a set of edges connecting these variables pairwise. Edges represent relations of possible manipulation. These possible manipulations impose a direction on the edges: If X and Y are two variables connected by an edge, the edge is directed toward Y if and only if there is a possible intervention on X, such that, if the intervention changed the value of X but the values of all other variables in the set V were held fixed at some value, the value of Y would undergo a change.

This framework allows the definition of various causal concepts. For the purposes of this paper, it will be sufficient to concentrate on the notion of a *direct cause*. "A necessary and sufficient condition for X to be a direct cause of Y with respect to some variable set V is that there be a possible intervention on X that will change Y (or the probability distribution of Y) when all other variables in V besides X and Y are held fixed at some value by interventions." (Woodward 2003, p. 55) A necessary condition for X to be a direct cause of Y, relative to a given set V of variables, is that V contains no variables that are causally intermediate between X and Y.

This definition makes crucial use of the notion of an intervention. Interventions are represented by exogenous variables, whose values are not determined by the values of the variables within V, but from outside the system. Typically – but not necessarily – the value of an intervention variable is determined by a human experimenter. Here is Woodward's (2003) definition of an intervention variable.

"I is an intervention variable for X with respect to Y if and only if I meets the following conditions:

I1. I causes X.

I2. I acts as a switch for all the other variables that cause X. That is, certain values of I are such that when I attains those values, X ceases to depend on the values of other variables that cause X and instead depends only on the value taken by I.

I3. Any directed path from I to Y goes through X. That is, I does not directly cause Y and is not a cause of any causes of Y that are distinct from X except, of course, for those causes of Y, if any, that are built into the I-X-Y connection itself; that is, except for (a) any causes of Y that are effects of X (i.e., variables that are causally between X and Y[1]) and (b) any causes of Y that are between I and X and have no effect on Y independently of X.

I4. I is (statistically) independent of any variable Z that causes Y and that is on a directed path that does not go through X. » (Woodward 2003, p. 98)

In Woodward's example mentioned above, the intervention variable represents the experimenter's decision to send a given child to private school. This is an intervention

[1] This clause does not apply to the case of direct causation.

according to the definition because 1) it causes P, in the sense that the intervention determines whether the child attends a private school. 2) It is part of the idea of a randomized experiment that only the experimenter's decision to attribute a given child to the experimental or control group determines whether she attends a private school or not. 3) The decision to put a child in the experimental group that attends a private school does not *directly* influence the child's scholastic achievements, i.e. if it influences them, it does so only by way of her attending a private school. 4) The very idea of a randomized experiment consists in making the determination of the value of P independent of all other variables, and in particular of variables that might influence A.

In the physical example mentioned above, raising the voltage across the copper wire satisfies these conditions on an intervention on the electric current P: 1) A change of the voltage I causes a change in the electric current P. 2) The electric current P is determined only by the voltage I (given the electrical resistance of the wire). 3) The voltage does not directly cause the wire to release heat but only through the flow of electrical current it causes. 4) The voltage is statistically independent of other causes of the wire's releasing heat, such as the room temperature, or various kinds of radiation.

2. Association Laws

Woodward's conditions are *not sufficient* for X being a direct cause of Y. My argument for this claim involves functional association laws, which are symmetric in the sense that they express *mutual* functional dependence between two variables X and Y, given other variables. Such laws create the following conceptual problem for Woodward's analysis of direct causation. Intervening on X (while holding other variables fixed) changes Y, so that X should be a direct cause of Y; but intervening on Y changes X, so that Y should be a direct cause of X. This mutual dependence holds for every particular system to which the law applies, at every instant. This can be made explicit by taking the relevant variables to be specific for the system *s* and the time *t*[2]. Let P be the generic variable representing electric current and A the generic variable representing heat release. The formalism of graphs is mostly used at this generic level: statistical tools are used to determine whether P is correlated with A. But causal processes take place at determinate places and times. Therefore, we shall consider specific variables: P(s,t) is the variable representing the electric current in wire s at time t. In general, let X(s,t) be the specific variable representing the value X takes in system s at t.

In terms of specific variables, the problem is this. If X(s,t) and Y(s,t) are related by an association law, the interventionist analysis yields the result that X(s, t) is a direct cause of Y(s, t), and that Y(s, t) is a direct cause of X(s, t). This is incompatible with the asymmetry of causation. I conclude that the relation characterized by Woodward's conditions cannot be causation. Rather, it is a relation that it not asymmetric and easily confounded with causation. I suggest it is nomological dependence.

It might seem that using specific variables implied changing the subject from type level causation, which is Woodward's topic, to token-level causation[3]. This would be a misunderstanding. The idea is rather that there is an important difference between the relations of variables at the type or generic level, which can be brought out in the clearest way by considering specific variables. The examination of the relations between these specific

[2] Spohn (2000 ; 2006) uses specific variables in this sense, whereas Woodward (2003) following Spirtes, Glymour and Scheines (2000) and Pearl (2000), uses generic variables.

[3] The causal relation between specific variables bears close resemblance to but should not be confused with (what Woodward and others call) "actual causation". Actual causation, as defined by Woodward (2003, p. 77) is a relation between specific *values* of variables, whereas specific variables are still variables.

variables shows that there are two types of dependencies between generic variables that are mixed together if one ignores the level of specific variables: relations of mutual dependence between generic variables on one hand and relations of non-mutual asymetric dependence between generic variables. The latter can be causal but not the former.

There is my argument. In order to evaluate type level-causal claims bearing on the relation between generic variables, it is necessary to study interventions. These interventions bear on particular systems and can be represented by specific variables. At the level of the particular interventions, there appears a difference between interventions in which the (specific) effect variable necessarily bears on a later moment that the (specific) cause variable and interventions in which it is possible that the effect variable bears on the same instant as the cause variable.

The law I will take as an example is a system law[4] valid for all devices containing rotating electrically charged masses. It says that the angular momentum due to the rotation of the mass is proportional to the magnetic moment due to the rotation of the electric charge. It can be easily derived from two laws of nature.

(1) The angular momentum **L** of a mass m rotating with speed **v** in a circle with diameter **r** is

$$\vec{L} = \vec{r} \times m\vec{v}$$

where **r** is the particle's position in a coordinate system centered at the centre of rotation, and × denotes the cross product. Variables in boldface are vectors.

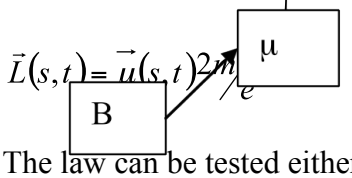
(2) The magnetic moment **μ** of an electric charge e rotating with speed **v** in a circle with radius **r** is

$$\vec{\mu} = {}^{1}\!/_{2}\, e\vec{r} \times \vec{v}\,.$$

A little calculation putting (1) and (2) together yields

$$\text{(MS)}\ \vec{L} = \vec{\mu}\, {}^{2m}\!/_{e}$$

This [v] tic stirrer law" (MS), according to which **L** and **μ** are functions of each other, is an a on la orresponding specific variables, characterizing a determinate magnetic stir determinate instant t, are **L**(s, t), **μ**(s, t). Each particular magnetic stirrer obeys at each moment t a specific law that is an instance of the general magnetic stirrer law.



The law can be tested either by manipulating **L**, e.g. by mechanically increasing the speed **v** of the rotating mass, while holding fixed the mass m and the charge e, and observing the change in **μ**, or by manipulating **μ**. The latter can be done, e.g., by increasing the strength of the magnetic field **B**, which accelerates the rotation of the charge. The law (MS) is then tested by observing whether and how much this influences **L**. Fig. 1 shows a graph representing variables **L** and **μ** connected by (MS), as well as two intervention variables: **v** represents the speed of the rotating object; **B** represents the magnetic field. The controversial question is whether the edge between **L** and **μ** can be given a causal interpretation.

[4] The domain of such a law is not universal, as is the case with general laws of nature, but consists of all systems of a given type. Cf. Schurz (2002). Cummins (2000) calls them "in situ" laws. Cartwright (1999) calls systems obeying such laws "nomological machines".

v = speed of the rotating object **L** = magnetic moment
B = magnetic field **μ** = angular momentum

Fig. 1. Graph representing the variables relevant for testing the magnetic stirrer law (MS)

Let us take the mass m and charge e of the rotating object to be fixed. An intervention on **L**(s,t) by manipulating the speed **v** will change **μ**(s,t) and an intervention on **μ**(s,t) by manipulating the strength of the magnetic field **B** will change **L**(s,t). Therefore, the interventionist model leads to the conclusion that **L**(s,t) and **μ**(s,t) are mutual causes of each other. It is crucial here to distinguish between generic and specific variables. We will return to generic variables later. Contrary to what the interventionist analysis implies, two specific variables, representing different properties of the same system at the same time, cannot be mutual causes of each other. Here is a simple argument for this claim. Let us assume the interventionist analysis. Given that manipulating **L**(s,t) changes **μ**(s,t) and manipulating **μ**(s,t) changes **L**(s,t), **L**(s,t) is cause of **μ**(s,t) and **μ**(s, t) is cause of **L**(s,t). Now suppose causation is transitive. Then we get the absurd consequence that **L**(s,t) and **μ**(s,t) are causes of themselves, because, e.g., **L**(s,t) causes **μ**(s,t) which causes **L**(s,t), so that **L**(s,t) causes **L**(s,t).

Now, one may object that this reasoning depends on the controversial thesis that causation is transitive. However, all counterexamples to the transitivity of causation that can be found in the literature belong to one of two categories; however, as I will show now, association laws belong to neither of these categories. In counterexamples of the first category, transitivity seems to be violated to the extent that it is left unspecified whether the items that are causally related are events or facts. The second category concerns cases where the cause or the effect (or both) is a *negative* fact; double prevention is a particular case[5].

The following case described by Ehring (1987) belongs to the first category. Smith puts potassium salts in the fireplace, making the fire in the fireplace purple. The fire then lights a log lying nearby. There is a causal chain from Smith's throwing potassium salts in the fireplace to the log's taking fire, but it seems wrong to say the former event is a cause of the latter[6]. One way to account for this scenario without abandoning the requirement of transitivity is to take the terms of causal relations to be facts (Kistler 2001) or aspects of events (Paul 2004). At the level of facts, there is no causal chain relating Smith's act to the log's inflammation, because there is no common middle term. The effect of the first causal process, the fire in the fireplace being purple, is not identical with the cause of the second

[5] See Bennett (1987), Hall (2004a).
[6] Other examples can be found in McDermott (1995), Hall (2000/2004b), Paul (2004).

process, the fire being hot. Without a causal chain, the question of transitivity does not even arise. There is an illusory appearance of a causal chain as long as one doesn't distinguish the fire's becoming purple from the fire's being hot.

The causal chain leading from $\mathbf{L}(s,t)$ to $\mu(s,t)$ and then again to $\mathbf{L}(s,t)$ does not belong to this category of counterexamples to transitivity. The appearance of a chain in the potassium salts case depends on the ambiguous specification of the middle term. Here, there is no such ambiguity. The middle term is exactly $\mu(s,t)$. Even if we consider $\mu(s,t)$ as a fact or an aspect of an event, we still get the result that $\mathbf{L}(s,t)$ (and $\mu(s,t)$ for that matter) causes itself.

The second category of counterexamples to transitivity involves causal relations where the cause or the effect is a negative fact. Let us consider a case of so-called double prevention that Hitchcock (2001) attributes to Ned Hall[7]. A hiker sees a rock falling, which makes her duck to avoid it. The fact that the hiker didn't get hurt by the rock makes her continue her trip. This is a case of double prevention, in the sense that the hiker's ducking prevents the falling of the rock from preventing her from continuing her trip. It seems wrong to say that the rock's falling causes the continuation of the trip, although there seems to be a causal chain linking the former, via the hiker's ducking, to the latter. As with the first category of counterexamples, it is possible to defend the transitivity of causation by denying that there is a causal chain. There are independent reasons for denying that negative facts, such as the fact that the hiker is not hit by the rock, can be causes or effects[8]. Negative facts enter into relations of explanation, which may be indirectly grounded on (and made true by) causal processes, but they are no terms of causal relations. However, the chain from $\mu(s,t)$ to $\mathbf{L}(s,t)$ and then back to $\mu(s,t)$ does not belong to this second category either, simply because $\mu(s,t)$ and $\mathbf{L}(s,t)$ are no negative facts.

This suggests[9] that the relations between $\mu(s,t)$ and $\mathbf{L}(s,t)$ do not belong to any category of relations that give rise to the illusion of a causal chain. But then there really is a chain of determination by which each of the variables $\mu(s,t)$ and $\mathbf{L}(s,t)$ indirectly determines itself via the other variable and two instances of (MS). However, it is absurd that a specific variable causes itself. I conclude that the relation of determination expressed by an association law such as (MS) is not causation.

Let us take stock. Our aim is to evaluate the interventionist conception of causation. Given that association laws guarantee manipulability, I have argued that the existence of a manipulability relation between two specific variables $\mu(s,t)$ and $\mathbf{L}(s,t)$, i.e. the fact that an intervention on one variable can be used to manipulate the other, is not sufficient for the existence of a causal relation between these variables. The argument proceeded by reductio, showing that if we suppose the transitivity of causation, the opposite hypothesis (i.e. that manipulability is sufficient for causation) leads to the absurd result that specific variables such as $\mathbf{L}(s,t)$ cause themselves.

3. Defending interventionism

Here is what appears to be a straightforward way of defending the hypothesis that manipulability is sufficient for causation, even if laws of simultaneous association are taken into account. One might argue that there can be no "causal loop" from $\mathbf{L}(s,t)$ to $\mu(s,t)$ and back to $\mathbf{L}(s,t)$ because it is impossible to intervene both on $\mathbf{L}(s,t)$ and on $\mu(s,t)$ at the same time t in the same system s. Indeed, given that these variables stand in a relation of mutual

[7] Hitchcock (2001, p. 276) indicates that it figures in an unpublished version of Hall (2004a).
[8] I have argued for this claim elsewhere (Kistler 2006).
[9] It does not establish it. Maybe the relations in the chain $\mathbf{L} - \mu - \mathbf{L}$ are not transitive for some other reason. But I can't think of any such reason.

functional dependence, one cannot independently fix both **L** and **μ** "from outside the system". However, in the framework of the interventionist account, in order to justify the claim that both **L(s,t)** causes **μ**(s,t) and **μ**(s,t) causes **L**(s,t), it is only required that for all s and t, it is possible to carry out *each* intervention. The defense of interventionism under consideration fails by committing a fallacy relative to the scope of the temporal quantifier "at every instant t". To justify that **μ**(s,t) causes **L**(s,t) and **L**(s,t) causes **μ**(s,t), it is necessary and sufficient that: for all s and all t, it is possible to manipulate **L**(s,t) by intervening on **μ**(s,t), and for all s and all t, it is possible to manipulate **μ**(s,t) by intervening on **L**(s,t). Given the mutual functional dependence of **L** and **μ**, this is equivalent to:

(s)(t)PM **L**(s,t) and (s)(t) PM **μ**(s,t),

Where PM represents the operator "it is possible to manipulate".

However, it is *not* necessary that: for all s and all t, it is possible to manipulate both **L**(s,t) by intervening on **μ**(s,t) and to manipulate **μ**(s,t) by intervening on **L**(s,t). Given the mutual functional dependence of **L** and μ, this is equivalent to:

(s)(t)(PM **L**(s,t) and PM **μ**(s,t))

Only the latter condition cannot be satisfied because one cannot perform both manipulations on the same system at the same moment.

A more radical move to defend interventionism would be to claim that the direction of causation is not objectively determined, but depends on perspective[10]. According to whether one *considers* an intervention on **L** or on **μ**, **μ** is cause of **L** or **L** is cause of **μ**. Then it seems to be enough not to take both perspectives at the same time to avoid the result that there is a causal loop at the level of specific variables. I take it to be unsatisfactory to consider the direction of causation as not objectively determinate. But this move cannot anyway really rescue interventionism. Let us assume that the direction of the causal relation between **L** and **μ** is determined by the direction a given cognitive agent *considers*. True, it is impossible to intervene on both **L** and **μ** in the same system at the same instant. However, nothing stands in the way of *considering* both interventions at the same time. First, there may be two agents each of whom considers one of the two directions. Second, a single agent may consider both directions at the same time. She can, e.g., draw two diagrams next to each other, one of which represents an intervention on **μ** and the other an intervention on **L**, and then consider both diagrams at the same time. Both possibilities show that it is possible to endorse both perspectives (the perspective according to which **L** causes μ, and the perspective according to which μ causes **L**) at the same time. Therefore, if the direction of a causal relation is determined by the perspective of a cognitive agent, causal relations in both directions can coexist at the same time. The refutation by reductio goes through as before: Given transitivity, each of the relata is its own cause.

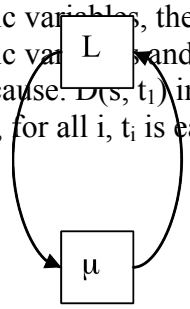
4. Mutual determination and feedback cycles

Graphs are mostly used to represent relations between *generic* variables. At the level of generic variables, feedback cycles are common, in which the causal influence of variable X on variable Y coexists with a reverse influence of variable Y on variable X[11]. However, such feedback cycles are very different from relations of mutual determination grounded on association laws. The difference appears clearly as soon as time is explicitly represented. Take the economic feedback circle in which the increase of demand D of a good increases its

[10] I criticize this move, suggested by Fair (1979), in Kistler (2006).
[11] Pearl (2000, p. 12ff.).

price P, which increase in turn lowers the demand D. The influences from D on P and from P on D are causal but not simultaneous.

At the level of generic variables, which are independent of the times of their instances, price and demand form a circle (fig. 2). As soon as we switch from generic variables to time-specific variables, the circle is replaced by a zig-zag line (fig. 3). Fig. 3 shows a graph with specific variables and a temporal dimension. It shows that effect is delayed with respect to its cause. $D(s, t_1)$ influences $P(s, t_2)$, which influences $D(s, t_3)$, which influences $P(s, t_4)$ etc, where, for all i, t_i is earlier than t_{i+1}.



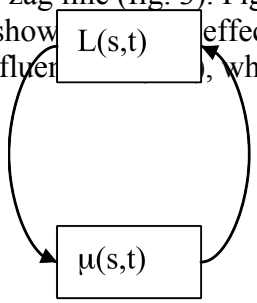


Fig. 2: Cycle involving generic variables

Fig. 3: Delayed influence between different time-specific variables

If we make the same move from generic to time-specific variables for variables related by a law of simultaneous association, the result is different.

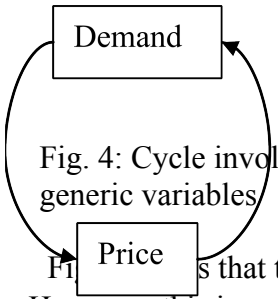


Fig. 4: Cycle involving generic variables

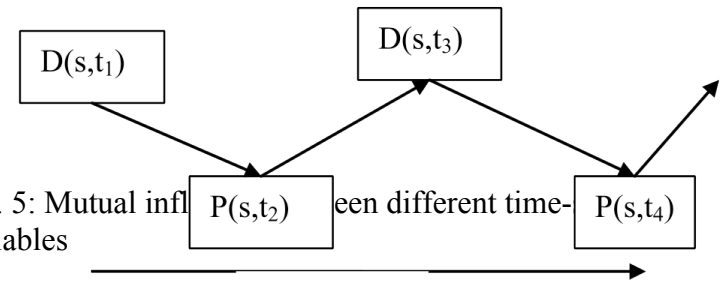


Fig. 5: Mutual influence between different time-specific variables

Figure 5 shows that the relation between the generic variables price and demand is no feedback cycle. However, this is not apparent at the level of generic variables (fig. 4), where their relation cannot be distinguished from a feedback cycle. Two variables form a feedback cycle if two conditions are satisfied. 1. The generic variables form a circle (as in fig. 2 and fig. 4). 2. Specific variables taken for the same system at the same time do *not* form a circle. According to these criteria, price and demand really form a feed-back cycle. In their case, specific variables for the same system at the same time are not related by a circle, and thus, the

problem of self-ca[illegible] t arise. [illegible] d μ do [illegible]lback cycle, because the relevant specific variables also form a circle.

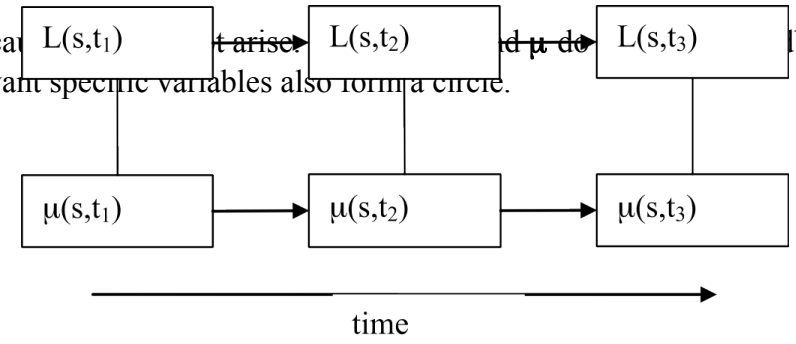


Fig. 6. Time evolution of specific variables linked by a simultaneous association law.

The difference between a simultaneous association law and a feedback law also appears clearly if we represent the evolution of the specific variables related by such laws in a graph with a temporal dimension (fig. 6). If, for every moment t_i, two specific variables $\mathbf{L}(s,t_i)$ and $\boldsymbol{\mu}(s,t_i)$ are related by a law of simultaneous association, at each moment t_i, the variables corresponding to that instant t_i stand in a relation of mutual dependence. These dependence relations appear in fig. 6 as vertical lines. If each of these dependence relations were causal, we would, given the transitivity of causation, get the absurd result that, at each moment, each of these specific variables was a cause of itself[12].

Could we not solve the problem raised by the mutual manipulability of variables linked by simultaneous association laws, simply by adding the requirement of a temporal delay to the condition of manipulability? $\mathbf{L}(s,t)$ does not cause $\boldsymbol{\mu}(s,t)$ and vice versa, although each can be used to manipulate the other, simply because these specific variables do not satisfy the requirement of temporal delay. In a way, this will be our conclusion. But it seems preferable that this requirement not be simply added ad hoc, but instead follow from more general considerations. Indeed, simply requiring that the effect follows the cause after some finite delay leaves it unclear why simultaneous causation does not exist, and why delayed manipulability reveals causation whereas simultaneous manipulability does not[13].

Spohn (2000; 2006) makes a move that looks similar but is in fact more radical. He requires that every specific variable represented in a model concerns a different time; in other words, an acceptable model must not contain more than one specific variable concerning a given instant. Contrary to the postulate we considered in the last paragraph, Spohn does not

[12] Maybe this distinction between genuine feedback cycles and mutually dependent variables lies behind Pearl's stipulation that "directed graphs may include directed cycles (e.g., X→Y, Y→X), representing mutual causation or feedback processes, but not self-loops (e.g., X→X)" (Pearl 2000, p. 12). At the level of generic variables, this stipulation seems completely unmotivated. Indeed, as Pearl explicitly proves (2000, p. 237), in the absence of other influences, transitivity holds: if X→Y and Y→Z (but no additional influence X→Z along any pathway independent from the pathway running through Y), then X→Z. Transitivity seems to imply directly that every "directed cycle" X→Y, Y→X necessarily entails the existence a "self-loop" X→X. One coherent interpretation of Pearl's remark would be to take "self-loops" to refer to what I have called relations of mutual dependence (which are circular both at the generic and at the specific levels), whereas directed cycles correspond to what I have called feedback cycles (which are circular only at the level of generic variables). However, Pearl's framework cannot give any justification for his exclusion of self-loops, which therefore seems ad hoc.

[13] We will see shortly (in section 5) that there are other reasons for which the requirement of temporal precedence is insufficient to save the interventionist analysis of causation, once association laws are taken into account.

stipulate that relations between simultaneous variables are never causal. Rather, he excludes the possibility to represent such simultaneous specific variables in the first place. This makes it trivially true that there are no causally related specific variables characterizing the same system at the same time. It is true simply because no model can contain such specific variables relative to the same system at the same time *at all*.

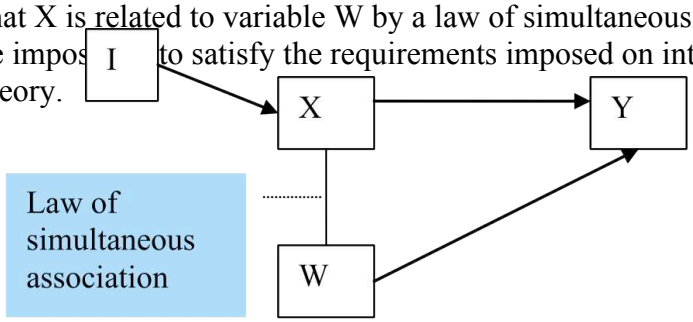
This requirement guarantees indeed that manipulability is sufficient for causation. It solves our problem. However, from the point of view of the representation of scientific methodology, the requirement of temporal precedence appears to be ad hoc and too strong.

1. First, it seems ad hoc to disallow the representation of relations between variables linked by an association law, insofar as the experimental investigation of their relation, leading to the discovery of that law, seems to follow exactly the same strategy as the discovery of causal relations: observation of statistical correlations and independencies on one hand, and experimental intervention on the other hand.
2. The second reason is even more important. The existence of a temporal delay can be the object of experimental enquiry. It may be a matter of empirical research whether a change in one of these variables leads to a simultaneous or a delayed change in the other variable. A model that excludes one of these possibilities by stipulation cannot make sense of such an investigation. Therefore, the requirement of temporal precedence imposes too strong limitations on the representation of the testing of scientific hypotheses.

5. Association laws and interventions

Let us then allow the representation of simultaneous specific variables in models of scientific enquiry. However, as we will see now, the simple presence of different specific variables characterizing the same system at the same time but linked by an association law raises a second problem for interventionism. There seem to be situations in which it is impossible to make interventions on such variables, according to Woodward's definition of interventions[14].

Let us suppose that X(s,t) and Y(s*,t*) are causally related variables which characterize different locations at different moments, so that causal influence spreads from s at t and reaches the point s* (at some distance from s) at some time t* (later than t). Let us further assume that X is related to variable W by a law of simultaneous association[15]. Then it will in general be impos to satisfy the requirements imposed on interventions by the interventionist theory.



[14] This is the reason – alluded to in the preceding note – for which the interventionist analysis cannot be saved by the requirement of temporal delay between the cause and effect variables.

[15] In general, X will be related by association laws to several variables W_1, W_2, … In what follows I will neglect this complication.

Fig. 7. Graph representing an intervention on X to test whether X causes Y, with X linked by an association law to W.

The problem is this. To be an intervention on X, I must not, beyond influencing X, also influence (clause I3) or be statistically dependent on (clause I4) variables that are on a causal path that leads to Y but does not lead through X. Let us assume there is an association law linking X to another variable W (fig. 7), and let us consider a type of situation in which the law applies without exception. Given the law linking X to W, it is impossible to intervene on X without at the same time intervening on W. By virtue of the law of functional dependence linking X and W, manipulating X by intervention I is necessarily (by nomological necessity) also manipulating W. This means that there cannot be a situation such as represented in fig. 7. If there is a law of functional dependence relating X and W, only fig. 8 can represent a situation in which X – and thus also W - is manipulated by I:

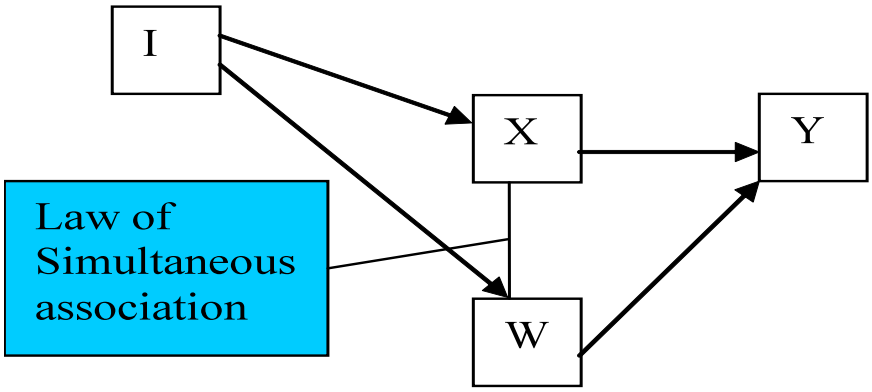


Fig. 8. Graph representing an intervention on X and W to test whether X causes Y.

Given that I can only manipulate X while also manipulating W, there cannot be a causal path from I to X without there also being a causal path from I to W. For the same reason, it is not justified to draw an arrow from X to Y without also drawing an arrow from W to Y: Changes in X are ipso facto (by nomological necessity) changes in W; thus, if a change in X is always followed by a change in Y, then a change in W (which goes, by nomological necessity, together with a change in X) is also always followed by a change in Y. In this case, conditions (I3) and (I4) cannot be satisfied: I is not an acceptable intervention on X according to (I3) if there is, parallel to the path from I to X, also a causal path from I to W, and then, parallel to the causal path from X to Y, also a parallel path from W to Y, as sketched in fig. 8. In such a situation, it is not true that “any directed path from I to Y goes through X”. I also violates condition (I4) because I is statistically correlated with W which causes Y and which is on a path to Y that does not go through X. Thus, it appears that satisfaction of the interventionist conditions is not necessary for causation. X may be a cause of Y although the conditions for an intervention on X cannot be satisfied.

Here is a situation of this type. Suppose we want to find out experimentally whether the magnetic moment (μ) of a magnetic stirrer causes electromagnetic radiation (R). We choose to intervene on μ via the magnetic field **B** (intervention variable).

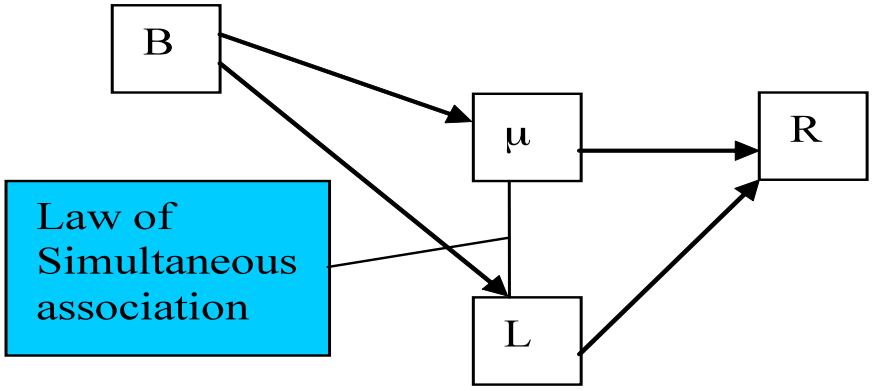


Fig. 9. Schema of an experiment to test the causal influence of magnetic moment (**μ**) on electromagnetic radiation (R)

If the mass m, charge e and radius of rotation of the stirrer are held fixed, and if (MS) $\vec{L} = \vec{\mu} 2m/e$

applies to the situation, it is impossible to make **μ** vary without making **L** vary as well. It follows that the intervention on **μ** violates both clause (I3) and clause (I4). Here is the reasoning that shows why for (I3). Given that, by virtue of the law (MS), any modification of **μ** goes by nomological necessity together with a modification of **L**, the reasons for taking B to be a variable allowing to manipulate **μ** are also reasons for taking B to be a variable allowing to manipulate **L**. Thus, it cannot be justified to draw an arrow from B to **μ** without also drawing an arrow from B to **L**. An analogous reasoning shows that it cannot be justified to draw an arrow from **μ** to R without also drawing an arrow from **L** to R. If any modification of **μ** is followed by a modification of R (which indicates a causal influence from **μ** to R), then it is also true by nomological necessity that any modification of **L** is followed by a modification of R. Thus, condition (I3) is violated: It is not true that "any directed path from **B** to R goes through **μ**", because there is a directed path from **B** to R that goes through **L** and not through **μ**. Interventions **B** on **μ** also violate (I4) because **B** is not statistically independent of **L**, which directly influences R on a path that does not go through **μ**. The same reasoning applies to interventions on **L**, which necessarily also influence **μ**. Therefore, according to the interventionist account, neither **μ** nor **L** are causes of R. This is certainly wrong: The rotation of electric charges causes electromagnetic radiation.

The problem I have just raised for magnetic moment and angular momentum is quite widespread. It arises not only for all causal claims bearing on variables that take part in laws of simultaneous association but also for all causal claims involving higher-level variables characterizing a complex system, in contexts in which the representation also contains lower-level variables on which the higher-level variables supervene. Furthermore, it arises for causal claims bearing on a determinable variable insofar as one of its determinables is also

represented. Before we get back to these other situations, let us look at some strategies for rescuing interventionism from this problem.

6 Can interventionism be saved?

All reactions to our problem that have appeared in the literature consider the specific case of causal claims involving a higher-level variable X, where the situation requires also the representation of lower-level variables W on which X supervenes. However, those reactions can quite easily be adapted to the case of association laws.

1. According to Woodward, if the appropriate interventions are impossible for "logical, conceptual, or perhaps metaphysical reasons, then that causal claim is itself illegitimate or ill-defined" (Woodward 2008, p. 224)[16]. We have seen that the appropriate interventions that might establish that X directly causes Y are impossible in case X is linked by a simultaneous association law to a variable W, which is linked to Y in the same way as X. Woodward's statement implies that such variables X cannot legitimately be judged to be causes. If there is a variable playing the role of W for every putative effect Y, X is epiphenomenal.
2. An alternative diagnosis is that interventionism is "inappropriate" (Woodward 2008, p. 256) to evaluate causal claims bearing on such variables. The interventionist framework is not applicable to causal relations where the cause variable X is linked by a simultaneous association law to another variable W where interventions on both X and W allow manipulating Y. Interventionism leaves us agnostic with respect to such variables: it neither justifies nor refutes the claim that X causes Y.

To avoid epiphenomenalism (1) or the result (2) that interventionism leaves us agnostic about the efficacy of variables X that are accompanied by "pre-empting variables" W, the interventionist framework may be modified in several ways so as to make it compatible with their efficacy. It would lead us too far to discuss all of them in depth. I will just mention a number of such possible modifications, and then consider the most promising a little more closely.

3. One may simply reject models containing variables standing in a relation of simultaneous nomic association or supervenience. By stipulation, an enquiry about a causal relation between X and Y must not include variables W that are lawfully linked to X or belong to X's supervenience base, and cause Y[17]. However, it seems arbitrary and ad hoc to exclude a scientific enquiry aimed at finding out which of variables X and W are causes of Y.
4. In case X and Y are higher-level variables, and W is in X's supervenience base, one may reject the causal closure of the lower-level variables in the supervenience basis of X and Y. Ultimately, this comes to challenging the completeness of physics. I will not go into this difficult topic here[18].
5. One may weaken the conditions imposed on interventions, so that it becomes possible to intervene on a variable X linked to another variable W by a simultaneous association law or by supervenience[19].

(I3) would, e.g., have to be modified in the following way (the modifications are in italics):

---

[16] This is also the conclusion of Shapiro and Sober (2007).

[17] Joseph Halpern (personal communication) has suggested this strategy. It is less restrictive than Spohn's mentioned above but still too restrictive.

[18] Kistler (2006) argues for this strategy.

[19] This strategy is suggested by Shapiro and Sober (2007): "To assess whether X causes Y, the common causes of X and Y must be held fixed, but not the microsupervenience base of X" (Shapiro and Sober 2007, p. 8).

(I3*) "Any directed path from I to Y goes through X *or through a variable linked to X by simultaneous nomic association, or through a variable in the supervenience base of X, or through a variable that is a determinate*[20] *of X*. That is, I does not directly cause Y and is not a cause of any causes of Y that are distinct from X except, of course, for those causes of Y, if any, that are built into the I-X-Y connection itself, *and except for variables linked to X by simultaneous nomic association, variables in the supervenience base of X, and variables that are determinates of X*; that is, except for (a) any causes of Y that are effects of X (i.e., variables that are causally between X and Y) and (b) any causes of Y that are between I and X and have no effect on Y independently of X, *and (c) any causes of Y that are linked to X by simultaneous nomic association, and (d) any causes of Y that belong to the supervenience basis of X, and (e) any causes of Y that are determinates of X*."

The problem with this strategy is that it seems to lead to systematic overdetermination[21]. Let us suppose that X and W are linked by a law of simultaneous nomic association, and let us suppose that X causally influences Y in that some modifications of X are systematically followed by a modification of Y. By virtue of the nomic association of X and W, necessarily, certain modifications of W are also followed by a modification of Y. Thus, W also causally influences Y. Y seems to be overdetermined because X's causal influence on Y is paralleled by the causal influence of W on Y. In the interventionist framework, there seems to be no way to single out one of the two nomically linked variables X and W as the cause of Y. An analogous reasoning shows that for every higher-level variable X that is a direct cause of some variable Y, there are lower-level variables belonging to the supervenience base of X which are also direct causes of Y. Shapiro and Sober (2007) accept this consequence. "If X causes Y, so does MSB(X)" (Shapiro and Sober 2007, p. 21; MSB is short for "micro-supervenience base"), so that "causation filters down" (p. 24) to lower levels. However, such generalized overdetermination seems to be incompatible with a basic idea of interventionism, which is that "genuinely competing or rival causal claims must make different predictions about what would happen under some possible intervention or interventions" (Woodward 2008, p. 224/5). The problem, from the interventionist perspective, is that the present strategy allows that there may be different claims of direct causation between variables (X causes Y, and W causes Y), which are justified by the same manipulation.

7. Difference Making

The most promising strategy of modifying interventionism is in terms of difference making (DM). DM seems to be able to justify at least in some cases that one and only one of a pair of non-causally related variables X and W is the cause of Y.

DM can justify the causal efficacy of variables within the interventionist framework, in two kinds of situation where the original interventionist analysis denies such efficacy:

1) The model of the situation contains variables standing in a relation of determinable to determinate, so that the efficacy of the former is threatened by the latter.

2) The model of the situation contains higher-level variables and lower-level variables in their supervenience base, so that each value of a higher-level variable corresponds to a set of values of a variable in its supervenience base.

However, as we will see in a moment, DM *cannot* justify the unique causal efficacy of

a) higher-level variables that are not multiply realized, and of

[20] A case of this sort will be considered in section 7.

[21] Several authors have suggested that overdetermination of this sort may be acceptable (Walden 2001, Sider 2003, Witmer 2003). Here, my point is that it seems difficult to accommodate it within an interventionist framework.

b) variables which are linked to their putative effect both directly and via variables associated to them by simultaneous association laws.

Menzies (2008) suggests spelling out difference making in terms of relations between values of variables. The idea is that specific values of the cause variable are correlated with specific values of the effect variable. "X=x rather than X=x* caused Y=y rather than Y=y*" (Menzies 2008, p. 207). In this condition, "causes" is understood in terms of counterfactual interventions, setting the value of the cause value respectively to x or x*.

Menzies shows that this account solves the problem of justifying, in the interventionist framework, the causal efficacy of determinable variables in models which also contain some of their determinates. The DM account delivers the intuitively correct result that a determinable variable and one of its determinates do not overdetermine their effect; furthermore, DM may justify the judgment that it is the determinable and not its determinate that causes a given effect. Menzies uses an example of Yablo's (1992) to illustrate his solution of the so-called "exclusion problem", according to which variables characterizing the brain systematically overdetermine the effects of variables characterizing the mental state of a cognitive system. According to Yablo, mental properties are determinables relative to brain properties which are their determinates. The following example illustrates how this model can justify the efficacy of a determinable property. Consider a pigeon conditioned to peck if and only if it perceives a red triangle. In a given situation, it perceives a crimson triangle and pecks. Taking perceiving a red target (X) and perceiving a crimson target (W) to be different variables, condition (I3) (and condition (I4)) on interventions cannot be satisfied with respect to the cause variable *X*. It is impossible to change the value of X without changing the value of W.

The strategy (5 in the last section) of weakening condition (I3) as in (I3*) leads to the incorrect result that the pigeon's pecking (Y) is overdetermined, insofar as both *X* and *W* appear as causes of Y. Interventions on *X* are possible according to (I3*), and some such interventions on X will lead to changes in Y. Some interventions on *W* (those which are also interventions on X) will also change the value of Y; so *W* also comes out as a cause of Y.

Intuitively, however, there seems to be a fact of the matter whether it is the fact that the pigeon perceives red or the fact that it perceives crimson that is the cause of its pecking. The difference making account reflects the experimental strategy suitable for finding this out: If the pigeon pecks at triangles that are red but not crimson, perceiving crimson is not the cause of its pecking. In that case, being crimson is not what makes the difference to whether it pecks or not. If the pigeon pecks at all red targets, but at no target of any other color, then perceiving red is what makes the difference to pecking. Thus, DM yields the intuitively correct result that the cause of the pigeon's pecking is one of the two variables *red* and *crimson*, but not both. The pecking is not overdetermined by them[22].

It seems possible to adapt the DM account to situations in which both the original interventionist account and its weakened version are unable to justify the efficacy of a higher-level variable X in causing a variable Y, because it *supervenes* on W which also seems to cause Y. Here is a sketch of such an adaptation. Let M be a (specific) mental variable representing a person's taking the decision to pick up a glass at time t. M takes the value 1 if

[22] The solution offered by difference making is equivalent to Yablo's (1992) solution in terms of "proportional" variables. If the pigeon pecks at any red target, not only at crimson targets, the pigeon's perceiving crimson is not "proportional" to the effect of pecking, which means that the perception of crimson is not necessary. Though devised to solve a different problem (late pre-emption), Lewis' (2000) concept of causation as "influence" provides an analogous solution. For X to influence Y, every modification in X must correspond to a modification in Y. In the case of Yablo's pigeon, the perception of crimson does not influence pecking, for switching from crimson to ruby is a modification in the alleged cause which is not accompanied by any modification in the effect, pecking.

and only it, at t, the person decides to pick up a glass[23]. Let B be the behavior of the person's picking up the glass. Let N be a lower-level variable describing a functional part of the person, i.e. her brain or some functional part of her brain that is essential in decision making. Each neurophysiologically different state of that part of the person's brain corresponds to a distinct value of N. Let us assume that M is multirealizable, so that there is a set of values of N $\{n_i, \ldots, n_k\}$ which are compatible with M=1.

Take a situation in which both M and B have value 1, and N takes value n_j. Which was the cause of B=1? $N=n_j$ or M=1? The original interventionist account cannot yield the result that M is a cause of B, because it is impossible to intervene on M without intervening on N. Supervenience implies that it is impossible to change the value of M without changing the value of N. The weakened account (strategy 5 in the last section) yields the result that both N and M are the cause of B. Both some interventions on the value of M, and some interventions on the value of N – i.e. those that change whether or not the value is in the set $\{n_i, \ldots, n_k\}$ - change the value of B. However, there seems to be a fact of the matter whether it is M=1 or $N=n_j$ that is the cause of B. One might conduct an experiment in the spirit of Mill's methods of agreement and difference, to find out whether one of them (and if yes, which) is both necessary and sufficient in the circumstances. It might turn out that $N=n_j$ is not necessary for B: in that case, as long as an intervention does not change the value of M (M=1), i.e. as long as it gives N a value belonging to the set $\{n_i, \ldots, n_k\}$, changing the value of N ($N=n_j$) does not change the value of B. On the other hand, it may turn out that M=1 is both necessary and sufficient for B=1, in the sense that setting M=1 guarantees that B=1 and setting M=0 guarantees that B=0.

However, DM cannot provide a solution to the problem of justifying causal claims bearing on variables that are connected to other variables by simultaneous association laws. Moreover, the justification of causal claims bearing on higher-level properties cannot be generalized to all cases either, for it depends on the multirealisability of the higher-level variable. Let me argue for the latter claim first: Take M to be a mental variable and N a variable describing the underlying brain state. Imagine it is discovered that M is not multirealised, so that M=1 is nomically correlated with a precise value of N, say N=1. Then M and N make exactly the same difference to B, because every intervention that changes the value of M changes the value of N and vice versa. DM cannot judge that one is the cause of B but not the other.

The same is true for our original problem case of variables linked by simultaneous association laws. Every intervention on $\mathbf{L}$ that changes $\mathbf{L}$'s value will necessarily also change the value of $\boldsymbol{\mu}$, as long as $\mathbf{r}$, m, and e are kept fixed. Equally, every intervention on $\boldsymbol{\mu}$ changes the value of $\mathbf{L}$. None of the two associated variables $\mathbf{L}$ and $\boldsymbol{\mu}$ is privileged over the other as the difference-maker for R (cf. fig. 9). Thus, the DM criterion cannot justify taking one rather than the other of them as the cause of R.

At this point, the defender of the DM account (and more generally the interventionist) may make one of two rejoinders. 1) She can simply accept overdetermination. It may sometimes happen that several variables are causes of a given variable. 2) She may argue that the fact that DM cannot decide which of two (or more) variables X and W causes a given variable Y shows that X and W are not really two different variables after all, but one variable with two names.

Indeed, in some situations, one of these replies is appropriate. However, this doesn't always seem to be the case. Let us start with the first reply. In the case of the magnetic stirrer, what seems correct to say is that the magnetic moment $\boldsymbol{\mu}$ and not the angular momentum $\mathbf{L}$

[23] Here I use the framework of "actual causation", where causes and effects are represented by values of variables, instead of variables themselves. Cf. Woodward (2003, p. 74-86).

causes electromagnetic radiation R. The reason is that there is a law of nature, according to which the rotation of an electric charge produces electromagnetic radiation, whereas there is no law according to which angular momentum produces electromagnetic radiation.

In the case of mental and neural variables describing the state of a cognitive system, the matter is more complex because there are less non-controversial psychological or psychophysical laws. However, for variables that figure in such laws, the reasoning is the same as for the magnetic stirrer law. The Rescorla-Wagner law (Rescorla and Wagner 1972) describes the variation of associative strength between a conditioned stimulus (CS) and an unconditioned stimulus as a function of a set of variables that can be interpreted as the surprise that the perception of the CS produces in an animal in a given situation. This law justifies taking the psychological variables composing that function (the surprise function) to be causes of associative strength, which is a psychological variable rather than an a variable describing the state of underlying neurons or synapses.

We have seen that the DM criterion is not always sufficient for the existence of causal influence. In cases where it is not, what appears to be required for the existence of causal influence between two variables is, in addition to DM, the existence of a law of nature linking cause and effect variables.

Let us look at the second reply. In some cases, it may indeed be reasonable to conclude that if it is impossible to distinguish by DM which of two variables X and W influences Y, then X=W. In such a case, there is after all just one variable with two labels. However, scientific reasons may plead against such a diagnosis. Laws are the crucial criterion for the identity and difference of variables. If there is a law in which X figures but no equivalent law in which W figures then X and W are not identical, although it may in a given situation be impossible to manipulate one without manipulating the other. In the stirrer example, it would definitely be wrong to identify **L** with **μ**, although these variables are, in the case of the stirrer, nomologically linked by the law (MS). The law does not say that **L** = **μ**, but that the values of the variables are related as $\vec{L} = \vec{\mu} 2m/e$. The variables **L** and **μ** are not identical simply because **L** figures in laws in which **μ** does not, and vice versa.

8. Conclusion

The problems that association laws raise for the Woodward's interventionist account of causation show that the satisfaction of interventionist criteria is neither sufficient nor necessary for the existence of a causal relation between two variables. The first problem was that the fact that X(s,t) makes a difference to Y(s,t) is not in itself sufficient for X(s,t) being a cause of Y(s,t). If s=s* and t=t, each of X(s,t) and Y(s,t) may make a difference to the other in case these variables are linked by a law of functional dependence. Then they cannot be causally related because the opposite hypothesis, together with the assumption of the transitivity of causation, leads to the absurd result that X(s,t) and Y(s,t) are both causes of themselves. The second problem was that the satisfaction of these conditions is not necessary either: We have seen that there are situations in which X(s,t) causes Y(s,t) but in which the conditions on an intervention on X(s,t) cannot be satisfied because X(s,t) is related to another variable W(s,t) by an association law.

The aim of this paper was mainly critical. However, let me end by suggesting that manipulability and its DM variant provide a satisfactory analysis of nomic dependence rather than of causation. A relation between variables X(s,t) and Y(s*,t*) that satisfies DM criteria can be either a relation of non-causal nomic dependence or a relation of causal dependence. If this is correct, it opens up the perspective of building upon DM to construct a new analysis of causation. Let me sketch one possible way in which this might be done. If t < t*,

manipulability and DM cannot be mutual. In general, the variable X(s,t) may characterize a system s, whereas variable Y(s*,t*) may characterize a different system s*. If t < t*, and if we assume that there is no backwards causation, it is not possible both that X(s,t) makes a difference to Y(s*,t*), and Y(s*,t*) makes a difference to X(s,t). Thus, if X(s, t) makes a difference to Y(s*, t*), this may be sufficient for the existence of a relation of *causal* nomic dependence between X(s, t) and Y(s*, t*). One way to analyse the relation of causal nomic dependence is in terms of nomic dependence and of a causal process. This is not the place to go into the analysis of the notion of a causal process. Let us suppose processes are grounded on the transmission of an amount of energy or some other conserved quantity[24]. Then X(s,t) can make a causal difference to Y(s*,t*) if there is a causal process extending from system s at t to system s* at t*. A special case is when s* is identical to s, so that the causal process consists in the evolution of one system, which is characterized by variables X and Y related by a functional association law.

In our example of the magnetic stirrer, let an intervention change the value of $\mathbf{L}$ at t, to set it to $\mathbf{L_1}$. The value of $\boldsymbol{\mu}$ changes simultaneously, becoming $\boldsymbol{\mu_1}$. Without further intervention, variables $\mathbf{L}$ and $\boldsymbol{\mu}$ still have the values $\mathbf{L_1}$ and $\boldsymbol{\mu_1}$ at a later time t*. The persistence of the rotating object from t to t* is a causal process. Therefore, one can justify the claim that setting $\mathbf{L}$ to $\mathbf{L_1}$ at t caused $\boldsymbol{\mu}$'s having the value $\boldsymbol{\mu_1}$ at t*, and also the claim that setting $\boldsymbol{\mu}$ to $\boldsymbol{\mu_1}$ at t makes a causal difference to $\mathbf{L}$'s having the value $\mathbf{L_1}$ at t*, in two steps.

1) One justifies the existence of a causal process linking the system at t to the system at t*. This may take the form of showing that the system at t is "genidentical"[25] to the system at t*. 2) One justifies the nomic dependence of $\boldsymbol{\mu}$ on $\mathbf{L}$, given that $\mathbf{r}$, m, and e are held fixed by the existence of the simultaneous association law (MS) relating these variables.

In this framework, it seems also possible to make sense of bottom-up and downward causation. If there is a relation of non-causal (simultaneous) bottom-up determination between a micro-variable A and a macro-variable B, intervening on A at t can cause B to have a certain value at t*. An example may be the determination of deficient behavior (B) of an animal that has been subject to "knocking out" of one of its genes (variable A). Knocking out the gene at t causes deficient behavior at t*, because 1) there is a causal process consisting of the animal persisting from t to t*, and 2) there is a relation of nomic dependence of B on A[26].

References

- Bennett, Jonathan (1987), Event Causation: the Counterfactual Analysis, *Philosophical Perspectives* 1, p. 367-368, repr. in Sosa et Tooley (1993), p. 217-233.
- Carnap, R. (1928), *Der logische Aufbau der Welt*, Hamburg, Felix Meiner, 1998.
- Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
- Collins, J., Hall, N. and Paul L. A. (eds.), Causation and Counterfactuals. Cambridge (Massachusetts), MIT Press, 2004.
- Cummins, R. (2000), How Does it Work? Vs. What Are the Laws? Two Conceptions of Psychological Explanation, in F. Keil & R. Wilson (eds.), *Explanation and Cognition*, Cambridge, MA: MIT Press, p. 117-45.
- Ehring, D. (1987), Causal Relata, *Synthese*, 73, p. 319–328.
- Fair D. (1979), Causation and the Flow of Energy, *Erkenntnis* 14, p. 219-250.

[24] Kistler (1998; 2006).

[25] Carnap (1928, § 159), Hawley (1999), Kistler (2001).

[26] I thank Olivier Darrigol, Isabelle Drouet and Gernot Kleiter, as well as my auditors in Toulouse, Liblice, Salzburg and Konstanz, where I have presented earlier versions of this paper, for many helpful suggestions.

- Hall, N. (2004a), Two Concepts of Causation, in Collins et al. (2004), p. 225-276.
- Hall, N. (2000/2004b), Causation and the Price of Transitivity, in Collins et al. (2004), p. 181-204.
- Hawley, K. (1999), Persistence and Non-supervenient Relations, *Mind* 108, pp. 53-67.
- Hitchcock, C. (2001), The Intransitivity of Causation Revealed in Equations and Graphs, *Journal of Philosophy* 98, p. 273-99.
- Kistler, M. (1998), Reducing Causality to Transmission, *Erkenntnis* 48, p. 1-24.
- Kistler, M. (2001), Causation as transference and responsibility, *in* Wolfgang Spohn, Marion Ledwig & Michael Esfeld (eds.), *Current Issues in Causation*, Paderborn, Mentis, p. 115-133.
- Kistler, M. (2001a), Le concept de génidentité chez Carnap et Russell, in: Sandra Laugier (ed.), *Carnap et la construction logique du monde*, Paris, Vrin, 2001, p. 163-188.
- Kistler, M. (2006), *Causation and Laws of Nature*, Londres, Routledge.
- Lewis, D. (2000), Causation as Influence, in Collins et al. (2004), p. 75-106.
- McDermott, M. (1995), Redundant Causation, *British Journal for the Philosophy of Science* 46, p. 523-544.
- Menzies, P. (2008), The Exclusion Problem, the Determination Relation, and Contrastive Causation, in J. Hohwy and J. Kallestrup, *Being Reduced*, Oxford, Oxford University Press, p. 196-217.
- Paul, L.A. (2004), Aspect Causation, in Collins et al. (2004), p. 205-224.
- Pearl J. (2000), *Causality: Models, Reasoning, and Inference*, Cambridge University Press.
- Rescorla, R.A. and Wagner, A.R. (1972) A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement, in A.H. Black and W.F. Prokasy, (eds.), *Classical Conditioning II*, p. 64–99. Appleton-Century-Crofts.
- Schurz, G. (2002), *Ceteris paribus* Laws: Classification and Deconstruction, *Erkenntnis* 57, p. 351-372.
- Shapiro, L. and Sober E. (2007), Epiphenomenalism. The Dos and Don'ts. In G. Wolter and P. Machamer (eds.), *Thinking about Causes: From Greek Philosophy to Modern Physics*, Pittsburgh, Pittsburgh University Press, p. 235-264.
- Sider, T. (2003), What's So Bad About Overdetermination? *Philosophy and Phenomenological Research* 67, p. 719-725.
- Spirtes, P., Glymour, C. and Scheines, R. (2000), *Causation, Prediction and Search*, Second edition, Cambridge (Mass.), MIT Press.
- Spohn, W. (2001a), Bayesian Nets Are All There is to Causal Dependence, in M.C. Galavotti et al (eds.), *Stochastic Causality*, Stanford: CSLI, and in W. Spohn, *Causation, Coherence, and Concepts: A Collection of Essay*s, Springer 2009.
- Spohn, W. (2001b), Deterministic Causation, *in* Wolfgang Spohn, Marion Ledwig & Michael Esfeld (eds.), *Current Issues in Causation*, Paderborn, Mentis, 2001, pp. 21-46.
- Spohn, W. (2006), Causation: An Alternative, *British Journal for the Philosophy of Science*, 57, p. 93-119.
- Walden, S. (2001), Kim's Causal Efficacy, *Southern Journal of Philosophy* 39, p. 441-460.
- Witmer, D.G. (2003), Functionalism and Causal Exclusion, *Pacific Philosophical Quarterly* 84, p. 198-214.
- Woodward, J. (2003), *Making Things Happen*, Oxford University Press.

- Woodward, J. (2008), Mental Causation and Neural Mechanism, in J. Hohwy and J. Kallestrup, *Being Reduced*, Oxford, Oxford University Press, p. 218-262.
- Yablo, S. (1992), Mental Causation, *Philosophical Review* 101, p. 245-280.